



Home



List

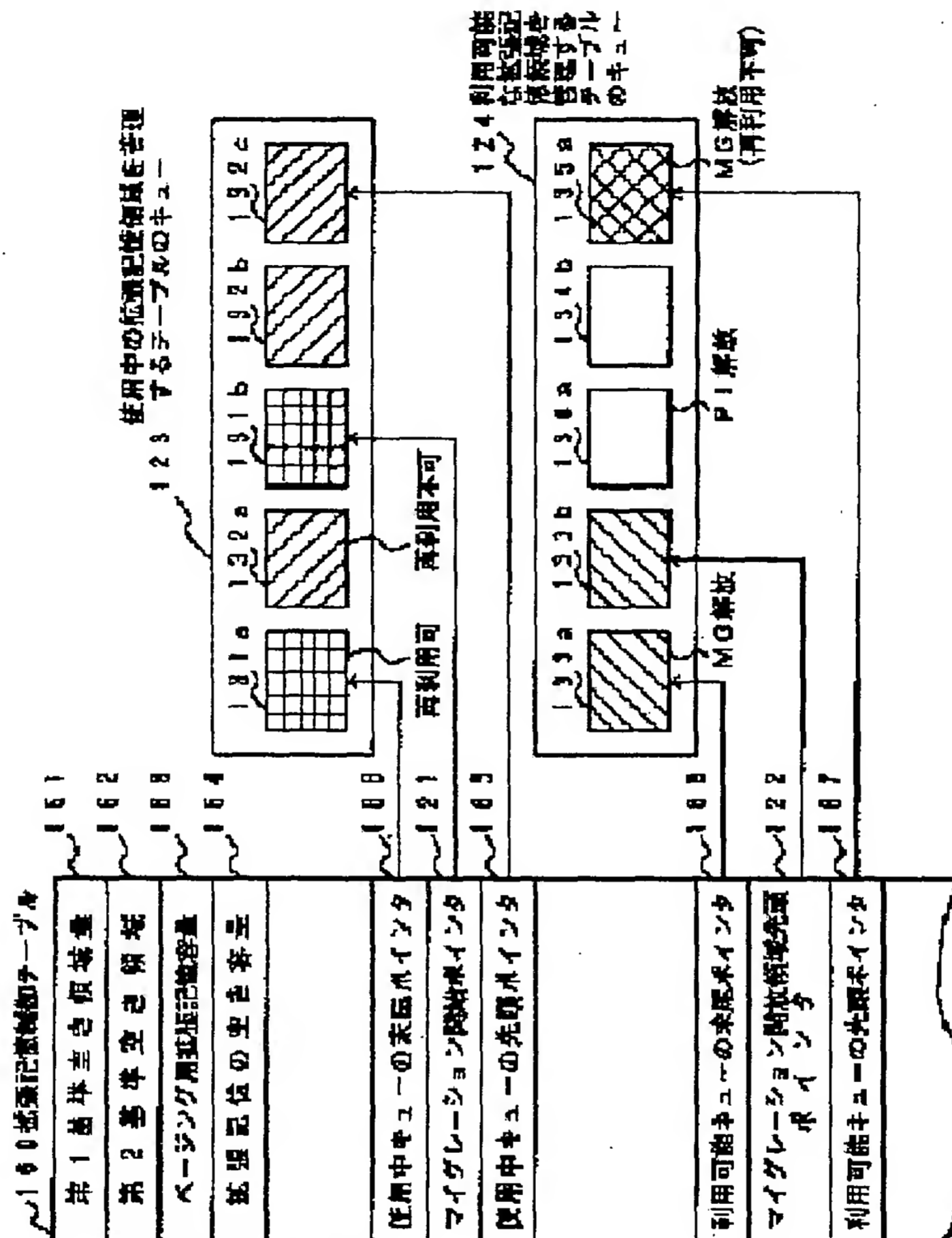
☐ Include

MicroPatent® PatSearch FullText: Record 1 of 1

Search scope: US Granted US Applications EP-A JP ; Full patent spec.

Years: 1991-2003

Patent/Publication No.: JP06059982



Order This Patent

Family Lookup

Find Similar

Legal Status

[Go to first matching text](#)

JP06059982 A

METHOD AND DEVICE FOR CONTROLLING VIRTUAL STORAGE
HITACHI LTD

Inventor(s): HOSOUCHI MASAOKI ; KATADA HISASHI ; ARAI TOSHIKI

Application No. 04212961 JP04212961 JP, Filed 19920810, A1 Published 19940304

Abstract: PURPOSE: To shorten page-out delay by purging the contents of a virtual storage from a high-speed external storage device to a low-speed external storage device in each regulated volume.

CONSTITUTION: When idle capacity 164 is less than the 1st reference idle capacity 161, only data corresponding to a capacity difference between the 1st level idle capacity and current idle capacity are transferred from an area 132e corresponding to the leading extended storage area management table entry (ESTE) of a using area queue 123 to a magnetic disk device through a main storage. When the capacity 164 is more than the capacity 161 and less than the 2nd reference idle capacity 162, fixed migration is executed after a fixed starting interval. The starting interval expected to be several seconds is set up by the CPU overhead of migration processing and the variation rate of the idle capacity and the capacity of migration based upon one starting is set up to a range capable of setting up the activity ratios of channels and a magnetic disk based upon paging I/O to less than

100% so as not to interrupt I/O to/from files.

Int'l Class: G06F01212; G06F01200 G06F01202 G06F01208

Patents Citing this One: No US, EP, or WO patents/search reports have cited this patent.



Home



List

For further information, please contact:
[Technical Support](#) | [Billing](#) | [Sales](#) | [General Information](#)

(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開平6-59982

(43)公開日 平成 6 年(1994) 3 月 4 日

(51)IntCl. ⁵	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F	12/12	Z	7608-5B	
	12/00	5 0 1 B	8526-5B	
	12/02	5 2 0	9366-5B	
	12/08	Y	7608-5B	

審査請求 未請求 請求項の数13(全 19 頁)

(21)出願番号	特願平4-212961	(71)出願人	000005108 株式会社日立製作所 東京都千代田区神田駿河台四丁目 6 番地
(22)出願日	平成 4 年(1992) 8 月10日	(72)発明者	細内 昌明 神奈川県川崎市麻生区王禅寺1099 株式会 社日立製作所システム開発研究所内
		(72)発明者	片田 久 神奈川県川崎市麻生区王禅寺1099 株式会 社日立製作所システム開発研究所内
		(72)発明者	新井 利明 神奈川県川崎市麻生区王禅寺1099 株式会 社日立製作所システム開発研究所内
		(74)代理人	弁理士 富田 和子

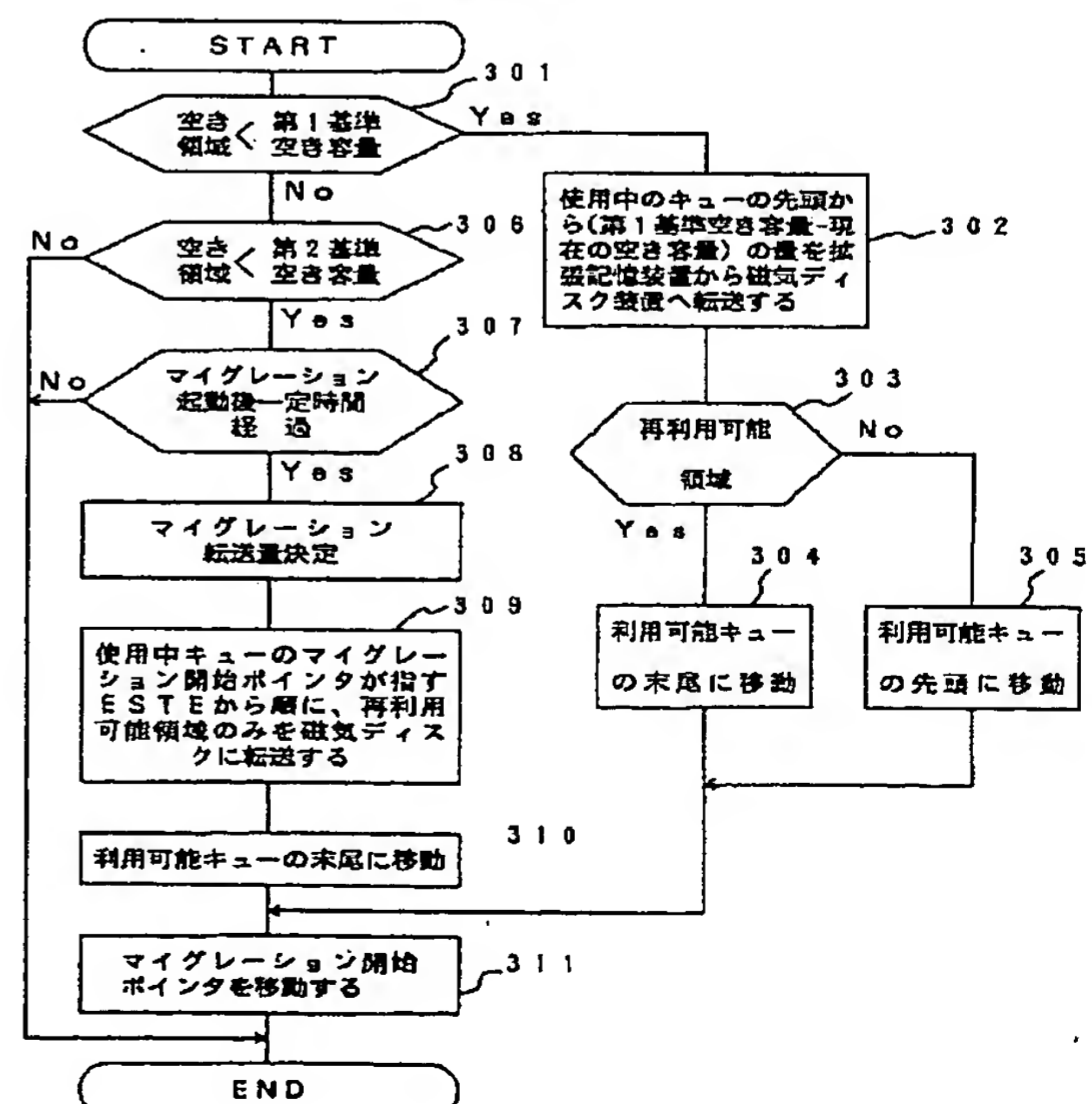
(54)【発明の名称】 仮想記憶制御方法および装置

(57)【要約】

【構成】 高速外部記憶装置から低速外部記憶装置への転送を開始する高速外部記憶装置の空き容量の基準値を2段階に設定し、空き容量に余裕のある場合でも、転送領域選択手段により再利用可能な高速外部記憶領域を選択し転送する。また、ページインにより利用可能となった高速外部記憶領域を低速外部記憶装置への転送により利用可能となった領域より優先して他の仮想記憶領域に割当ててゐる。

【効果】 高速外部記憶装置の利用率を落とさずに高負荷時の高速外部記憶装置の空き領域の供給不足を解消するとともに、低速外部記憶装置からの転送を削減することができ、仮想記憶制御方式の計算機システムの性能を向上させることができる。

マイグレーション処理 (図 8)



【特許請求の範囲】

【請求項 1】主記憶装置と、低速外部記憶装置と、高速外部記憶装置とを有する計算機システムにおける仮想記憶制御方法であって、

高速外部記憶装置の空き領域を検査し、
該空き容量が、予め定めた第 1 基準量より小さければ該第 1 基準量以上の空き容量が生じるよう高速外部記憶装置から前記低速外部記憶装置へ仮想記憶領域の内容を転送し、

前記空き容量が、前記第 1 基準量より大きく前記第 1 基準量より大きい予め定めた第 2 の基準量より小さければ、高速外部記憶装置から前記低速外部記憶装置へ仮想記憶領域の内容を規定量転送することを特徴とする仮想記憶制御方法。

【請求項 2】前記規定量は、前記空き容量に応じて可変としたことを特徴とする請求項 1 記載の仮想記憶制御方法。

【請求項 3】高速外部記憶装置の空き容量の前記第 1 基準量についての検査は、少なくとも主記憶装置から高速外部記憶装置への仮想記憶領域の内容の転送後に起動し、前記第 2 基準量についての検査は、少なくとも予め定められた一定時間間隔毎に起動することを特徴とする請求項 1 記載の仮想記憶制御方法。

【請求項 4】前記第 1 基準量より大きく前記第 1 基準量より大きい予め定めた第 2 の基準量より小さいときに行う転送の対象となる高速外部記憶装置の領域として、再利用可能な領域を選択することを特徴とする請求項 1 記載の仮想記憶制御方法。

【請求項 5】主記憶装置と、高速外部記憶装置および低速外部記憶装置を有する計算機システムの仮想記憶制御装置において、

前記主記憶装置上のデータを前記高速外部記憶装置に転送する第 1 の転送手段と、

前記高速外部記憶装置の空き容量があらかじめ定められた第一の基準容量以下の場合にあらかじめ定められた第一の量の前記高速外部記憶装置の内容を前記低速外部記憶装置へ転送する第 2 の転送手段と、

前記高速外部記憶装置の空き容量が前記第一の基準容量以上かつあらかじめ定められた第二の基準容量以下の場合にあらかじめ定められた第二の量の前記高速外部記憶装置の内容をあらかじめ定められた時間間隔において前記高速外部記憶装置から前記低速外部記憶装置へ転送する第 3 の転送手段とを有することを特徴とする仮想記憶制御装置。

【請求項 6】前記低速外部記憶装置へ転送する前記高速外部記憶装置の内容の前記第 2 の量を、前記高速外部記憶装置の空き容量に応じて変更する転送量調整手段を有することを特徴とする請求項 5 記載の仮想記憶制御装置。

【請求項 7】前記高速外部記憶装置として特定命令によ

りアクセス可能な半導体記憶装置である拡張記憶装置を用いるとともに、前記低速外部記憶装置として磁気ディスク装置を用いる請求項 5 記載の仮想記憶制御装置。

【請求項 8】各仮想記憶領域が前記主記憶装置、前記高速外部記憶装置および前記低速外部記憶装置のいずれの記憶装置のいずれの領域に割り当てられているかを管理するアドレス変換テーブルと、前記高速外部記憶装置の各領域がどの仮想記憶領域に割り当てられているかを示す割り当て情報を各エントリに格納して管理する高速外部記憶領域管理テーブルとを有し、各領域が再利用可能か否かを表わす再利用可能フラグを前記高速外部記憶領域管理テーブルに設け、前記第 1 の転送手段が転送を行う際、当該再利用可能フラグを設定し、前記第 3 の転送手段が転送を行う際、前記再利用可能フラグを参照して再利用可能な使用中領域を選択して当該データを前記低速外部記憶装置へ転送することを特徴とする請求項 5 記載の仮想記憶制御装置。

【請求項 9】前記高速外部記憶装置の各領域が再利用可能か否かは、前記アドレス変換テーブルにおいて、前記第 2 または第 3 の転送手段が前記高速外部記憶装置から前記低速外部記憶装置への転送を行う際に仮想記憶領域の高速外部記憶領域への割り当て情報が消失するか否かにより決定されることを特徴とする請求項 8 記載の仮想記憶制御装置。

【請求項 10】前記高速外部記憶領域管理テーブルの使用領域のエントリを連結した使用中キューを構成する手段と、該使用中キューの先頭を指示する第 1 ポインタ手段と、前記使用中キューの末尾を指示する第 2 ポインタ手段と、前記第 3 の転送手段が再利用可能か否かの検査を開始するエントリを指示する第 3 ポインタ手段とを有し、前記第 1 転送手段による転送を受けた高速外部記憶領域のエントリは前記使用中キューの末尾に接続されるとともに、前記第 2 の転送手段による転送の対象として前記使用中キューの先頭のエントリから順に選択されることを特徴とする請求項 8 記載の仮想記憶制御装置。

【請求項 11】前記高速外部記憶装置上のデータを前記主記憶装置へ転送する第 4 の転送手段と、前記高速外部記憶領域管理テーブルの利用可能領域のエントリを連結した利用可能キューを構成する手段と、該利用可能キューの先頭を指示する第 4 ポインタ手段と、前記利用可能キューの末尾を指示する第 5 ポインタ手段と、前記第 3 の転送手段が転送を開始する高速外部記憶領域のエントリを指示する第 6 ポインタ手段とを有し、前記第 3 の転送手段により転送された高速外部記憶領域のエントリは前記利用可能キューの末尾に接続され、前記第 4 の転送手段が転送した高速外部記憶領域のエントリは前記利用可能キューの第 6 ポインタ手段が指示するエントリの直前に挿入され、前記第 1 の転送手段による転送先の高速外部記憶領域としては前記利用可能キューの先頭のエントリの高速外部記憶領域が割り当てられることを特徴と

する請求項 8 または 10 記載の仮想記憶制御装置。

【請求項 12】主記憶装置と、高速外部記憶装置および低速外部記憶装置を有する計算機システムの仮想記憶制御装置において、

前記主記憶装置へ転送した仮想記憶領域の内容を保持していた前記高速外部記憶装置の領域を、前記高速外部記憶装置から前記低速外部記憶装置へ転送した仮想記憶領域の内容を保持していた前記高速外部記憶装置の領域よりも優先して、他の仮想記憶領域に割当てて手段を有する仮想記憶制御装置。

【請求項 13】前記高速外部記憶装置と前記低速外部記憶装置とを仮想記憶の記憶媒体とし、前記主記憶領域が不足した場合に前記主記憶装置から前記高速外部記憶装置へデータを転送するスワッピング手段と、前記高速外部記憶装置から前記低速外部記憶装置へのデータ転送手段とを有する計算機システムの仮想記憶制御装置において、

スワップアウトの原因を判別する手段を有し、前記高速外部記憶装置の容量が予め定められた容量以下の場合、前記判別手段により端末の入力待ちを起因としてスワップアウトされたと判断されたページは前記高速外部記憶装置に出力し、前記判別手段により主記憶容量不足を起因としてスワップアウトが必要と判断されたページは前記低速外部記憶装置へ出力する手段を有することを特徴とする仮想記憶制御装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、速度の異なる 2 つの外部記憶装置を備える計算機システムにおける仮想記憶制御装置に関するものである。

【0002】

【従来の技術】近年の仮想記憶方式の計算機システムでは、外部記憶装置として、磁気ディスク装置の他に磁気ディスク装置より高速の拡張記憶装置を備えたものがある。図 2 に、従来の拡張記憶装置を備えた仮想記憶制御装置の構成を示す。

【0003】図 2 において、101 は主記憶装置、102 は拡張記憶装置、103 は磁気ディスク装置、104 は仮想記憶、105 は仮想記憶制御装置である。仮想記憶制御装置 105 は、外部記憶装置（拡張記憶装置 102 または磁気ディスク装置 103）から主記憶装置 101 へ仮想記憶の内容をページ単位に転送するページイン制御部 111、主記憶装置 101 から外部記憶装置へ仮想記憶 104 の内容をページ単位に転送するページアウト制御部 112、主記憶装置 101 と外部記憶装置との間で仮想記憶 104 の内容を空間単位に転送するスワッピング制御部 113、拡張記憶装置 102 から磁気ディスク装置 103 に仮想記憶 104 の内容を転送するマイグレーション制御部 114、拡張記憶の領域を管理する拡張記憶領域管理部 115 から構成される。ここで、仮

想記憶の内容をページ単位に外部記憶装置から主記憶装置 101 へ取り込むことをページインといい、逆に主記憶装置 101 から外部記憶装置へ追い出すことをページアウトという。また、仮想記憶の内容を拡張記憶装置 102 から磁気ディスク装置 103 へ追い出すことをマイグレーションという。

【0004】拡張記憶装置 102 を備えた計算機システムにおける仮想記憶制御方式は、例えば特開昭 63-20552 号号に記載されている。

【0005】図 3 に、従来のマイグレーション方式のフローチャートを示す。マイグレーション制御部 114 では、拡張記憶装置 102 の空き領域を仮想記憶のページに割当てた際に（ステップ 201）、拡張記憶装置 102 の空き容量を調べ（ステップ 202）、空き容量が一定量以下になると、割当て済の拡張記憶装置 102 の領域の一部を主記憶装置 101 の一部を介して磁気ディスク装置 103 に転送保存し、拡張記憶装置 102 の空き領域を一定量以上に保つ（ステップ 203）。ページアウト制御部 112 では、主記憶装置 101 からあるページを追いつくそうとすると、拡張記憶装置 102 に空き領域があれば、当該ページにその空き領域を割当て、主記憶装置 101 の当該ページ内容を拡張記憶装置 102 へ出力する。拡張記憶装置 102 に空き領域がない場合は磁気ディスク装置 103 へ出力する。ページイン制御部 111 では、主記憶装置 101 に存在しない内容をページインしようとするとき、拡張記憶装置 102 を調べ、そこにページインしようとするページ内容が存在すれば、拡張記憶装置 102 から当該ページ内容を主記憶装置 101 へデータを転送する。

【0006】特開昭 63-128454 号公報には、拡張記憶領域管理部 115 の構造が記載されている。拡張記憶装置 102 の各ページを管理する拡張記憶領域管理テーブルは、使用中のテーブルエントリ 131、132 のキュー 123 と、利用可能（空き）のテーブルエントリ 133～135 のキュー 124 とにキューイングされる。拡張記憶領域の割当時には、利用可能なエントリのキュー 124 からはその先頭（図で右側）のエントリ 135a が選択され、使用中のエントリのキュー 123 の末尾（図で左側）に移される。また、マイグレーション時には、使用中のキュー 123 の先頭のエントリ 132c が選択されて開放され、利用可能なキュー 124 の末尾に移される。

【0007】また、「拡張記憶を利用した仮想記憶制御方式並びに性能評価」（情報処理学会 OS 研究会，1991. 3. 15）には、拡張記憶装置の実効容量拡大のため、ページイン時に拡張記憶装置の領域を解放する OS 重複ページ消去機能が記載されている。

【0008】さらに、HITAC マニュアル「センタ運営」（p131, 6180-3-101-10）には、主記憶容量不足を起因としてスワップアウトされた空間

は、時間が経過するほどスワップ勧告値が大きくなりスワップインされやすくなるが、端末入力待ちを起因としてスワップアウトされた空間は、一般に時間が経過するほどスワップインされる確率は低下する性質を持つことが記載されている。なお、スワップアウトとは、多重仮想空間方式において、1つの空間に割り当てられている主記憶上のすべてのページを外部記憶装置へ追い出すことをいい、スワップインとは、この逆の動作をいう。

【0009】

【発明が解決しようとする課題】従来のマイグレーション方式では、拡張記憶装置の空き領域が一定量以下になるとマイグレーションが行われるが、マイグレーションを開始する空き領域の量を少なく設定すると、負荷の変動が大きい場合、高負荷時にページアウトの多発により拡張記憶装置の空き領域がすぐに枯渇し、空き領域割当て待ちまたは磁気ディスクへのページアウトが発生してしまい、ページアウト遅延によりメモリ割当て待ち時間が増大する。その結果、処理性能が低下する。逆に、マイグレーションを開始する空き領域の量を多く設定すると、再利用不可能な拡張記憶装置の領域について磁気ディスク装置から主記憶装置へのページインが多発し、ページイン時間が増大する。また、高負荷時の性能を向上させるために拡張記憶の容量を増設すると、低負荷時に利用可能領域が増大し、拡張記憶装置の有効利用が図れなくなってしまう。

【0010】また、ページインにより解放された拡張記憶装置の領域が他のページに割当てられた場合には特に問題はないが、マイグレーションにより解放された領域が他のページに割当てられた場合では、マイグレーションにより解放された領域のデータをページインするために磁気ディスク装置へのI/Oが発生し、性能低下をもたらすという問題があった。

【0011】多重仮想記憶方式において、端末入力待ちを起因としたスワップアウトでは時間が経過するほどスワップインされる確率は低くなるが、主記憶容量不足を起因としたスワップアウトでは時間が経過するほどスワップインされやすくなる。従って、両者に同一のマイグレーションする拡張記憶装置の領域の選択アルゴリズムを適用すると、その一方について拡張記憶装置からスワップインされる割合が低下してしまう。

【0012】本発明の目的は、拡張記憶装置の利用率を落とさずに高負荷時の性能を向上させ、負荷の変動に対して安定した性能を維持する仮想記憶制御装置を提供することにある。

【0013】

【課題を解決するための手段】上記目的達成のため、本発明の仮想記憶制御方法は、主記憶装置と、低速外部記憶装置と、高速外部記憶装置とを有する計算機システムにおける仮想記憶制御方法であって、高速外部記憶装置の空き領域を検査し、該空き容量が、予め定めた第1基

準量より小さければ該第1基準量以上の空き容量が生じるよう高速外部記憶装置から前記低速外部記憶装置へ仮想記憶領域の内容を転送し、前記空き容量が、前記第1基準量より大きく前記第1基準量より大きい予め定めた第2の基準量より小さければ、高速外部記憶装置から前記低速外部記憶装置へ仮想記憶領域の内容を予め定めた量転送することを特徴とする。

【0014】また、本発明の仮想記憶制御装置は、主記憶装置と、高速外部記憶装置および低速外部記憶装置を有する計算機システムの仮想記憶制御装置において、前記主記憶装置上のデータを前記高速外部記憶装置に転送する第1の転送手段と、前記高速外部記憶装置の空き容量があらかじめ定められた第一の基準容量以下の場合にあらかじめ定められた第一の量の前記高速外部記憶装置の内容を前記低速外部記憶装置へ転送する第2の転送手段と、前記高速外部記憶装置の空き容量が前記第一の基準容量以上かつあらかじめ定められた第二の基準容量以下の場合にあらかじめ定められた第二の量の前記高速外部記憶装置の内容をあらかじめ定められた時間間隔をおいて前記高速外部記憶装置から前記低速外部記憶装置へ転送する第3の転送手段とを有することを特徴とする。

【0015】本発明の仮想記憶制御方法および装置の他の構成は、後述する実施例において詳述する。

【0016】

【作用】本発明による仮想記憶制御では、高速外部記憶装置の空き領域が一定量以上の場合でも規定量ずつ高速外部記憶装置から低速外部記憶装置へ仮想記憶の内容を追いつく。これによって、負荷の変動が激しくページアウトが集中した場合でも、あらかじめ負荷の低いときにマイグレーション（高速外部記憶装置から低速外部記憶装置へ転送）しておいた領域を割当てることにより、ページアウト遅延を短縮することができる。また、高速外部記憶装置の容量の増設と比較すると、同程度のページアウト遅延で、低負荷時における高速外部記憶装置の使用率を向上させることができる。

【0017】また、高速外部記憶領域の内容がマイグレーションにより低速外部記憶装置へ転送された後でもアドレス変換テーブルに高速外部記憶装置のアドレスが残っていれば、その領域が他の仮想記憶領域へ割当てられるまでは、アドレス変換テーブルに残っている高速外部記憶装置の同じ領域を再び割当てることにより再利用が可能となる。従って、アドレス変換テーブルの高速外部記憶割り当て情報のフィールドが他の用途に利用されているか否かの再利用可能フラグを、高速外部記憶管理テーブルに付加することにより、マイグレーションするときに、高速外部記憶領域管理テーブルを参照するだけで再利用可能か否かを判別できる。

【0018】さらに、第2基準容量に基づくマイグレーション時に、再利用可能な高速外部記憶装置の領域を選択して低速外部記憶装置へマイグレーションすることに

より、マイグレーションを開始する空き容量を高く設定しても、ページインの際に再利用の確率が向上し、低速外部記憶装置からのページインを増加を抑止することができる。

【0019】また、高速外部記憶装置から主記憶装置へのページインでは、高速外部記憶装置の領域が解放された他の仮想記憶領域に割当てられても、主記憶装置上に当該仮想記憶領域の内容が存在するので、当該仮想記憶領域のページアウトのときは高速外部記憶装置の別の領域を割当てればI/Oは発生しない。これに対し、マイグレーションされた高速外部記憶装置の領域が他の仮想記憶に割当てられてしまうと、低速外部記憶装置からページインする必要がある。従って、本発明では、ページインにより解放された高速外部記憶装置の領域を、高速外部記憶装置から低速外部記憶装置へ転送したため利用可能となった高速外部記憶装置の領域よりも優先して仮想記憶領域に割当てる。これにより、マイグレーションを開始する空き容量に関係なく、I/Oを減少させ、平均ページイン時間を短縮することができる。

【0020】また、高速外部記憶領域管理テーブルの使用中のエントリのキューに対して、マイグレーションを開始する位置を指示する第3ポインタ手段により、マイグレーションの判断が終了した再利用不可能な領域を管理するテーブルエントリと、マイグレーションの判断をしていない領域を管理するテーブルエントリとを同一のキューに配置することができ、マイグレーションのCPUオーバーヘッドを削減することができる。

【0021】さらに、高速外部記憶領域管理テーブルの利用可能なエントリのキューに対して、マイグレーションによって利用可能となったエントリの先頭位置を指示する第6ポインタ手段により、ページインにより利用可能となった高速外部記憶領域を、高速外部記憶装置から低速外部記憶装置へ転送したため利用可能となった高速外部記憶装置領域よりも優先して割当てる処理を、利用可能キューへのエントリの加入時の操作により容易に実現することが実現でき、割当て処理のCPUオーバーヘッドを削減することができる。

【0022】また、高速外部記憶装置の容量が小さいときに、高速外部記憶装置からページインする可能性の低い主記憶容量不足を起因としたスワップアウトを低速外部記憶装置へ出力することにより、端末の入力待ちを起因としたスワップアウトされたページの高速外部記憶装置上に存在する時間を長くすることができ、高速外部記憶装置から転送する割合を向上することができる。

【0023】

【実施例】以下、この発明の一実施例を説明する。

【0024】図1は本発明の仮想記憶制御システムを示すブロック図である。仮想記憶装置104の各領域（ページ）は、主記憶装置101、特定命令によりアクセス可能な高速外部記憶装置（半導体記憶装置）である拡張

記憶装置102、拡張記憶装置102に比べ低速な外部記憶装置である磁気ディスク装置103のいずれか一つまたは複数の装置に割当てられる。割当てる装置の選択制御は、仮想記憶制御装置105にて行う。

05 【0025】仮想記憶制御装置105は、ページイン制御部111、ページアウト制御部112、スワッピング制御部113、マイグレーション制御部114、拡張記憶領域管理部115、により構成される。スワッピング制御部113は、スワップ原因判別部116を含む。また、10 マイグレーション制御部114は、拡張記憶装置102から磁気ディスク装置103への転送量を決定する転送量調整部117と、拡張記憶装置102のうち磁気ディスク103へ追い出す領域を選択する転送領域選択部118とを含む。

15 【0026】拡張記憶装置102は一定の大きさの領域に分割され、各領域と一対一に対応付けられた拡張記憶領域管理テーブルエントリ（ESTE）により管理される。ESTE131～135は、後述するように拡張記憶領域管理部115において、使用中の拡張記憶領域の20 キュー123または利用可能（空き）拡張記憶領域のキュー124に接続され管理される。ここで、使用中のキュー123は、ページの内容により複数のキューに分割しても良い。

25 【0027】拡張記憶領域管理部115において、131aおよび131bは再利用可能領域のESTE、132aおよび132bは再利用不可能領域のESTE、133aおよび133bはマイグレーションにより解放された再利用可能領域のESTE、134aおよび134bはページインにより解放された領域のESTE、135aはマイグレーションにより解放された再利用不可能領域のESTEである。また、マイグレーションの開始位置を示すポインタ121と、マイグレーションにより解放された再利用可能領域のキューの先頭のESTEを示すポインタ122を有する。拡張記憶領域が再利用35 可能か否かについては、後に詳述する。

【0028】本発明は、拡張記憶装置102の空き容量に応じて拡張記憶装置102から磁気ディスク装置103への転送量を決定する点と、再利用可能な拡張記憶装置の領域と再利用不可能な領域を同一の使用キューに40 配置しながら両者を区別可能にした点、マイグレーションにより解放された領域とページインにより解放された領域を同一の利用可能キューに配置しながら両者を区別可能にした点、およびスワップ原因に応じて転送先を変更可能とした点とに特徴がある。

45 【0029】図4に、図1のハードウェアイメージを示す。仮想記憶制御装置105は、主記憶装置101上に存在する。図3において、主記憶装置101はCPU(Central Processing Unit)106の外部に接続されているが、内蔵していてもよい。拡張記憶装置102および50 磁気ディスク装置103の領域は、ページ単位の大さ

に分割され、主記憶に配置しきれないデータを格納する仮想記憶の記憶媒体として利用される。主記憶装置101と拡張記憶装置102との間はCPU106を介して、拡張記憶装置102と磁気ディスク装置103との間はCPU106およびIOP (Input Output Processor) 107および主記憶装置101またはCPU106およびIOP107を介して転送する。

【0030】仮想記憶装置104の各領域は、アドレス変換テーブル141のエントリと1対1に対応している。主記憶装置101上のデータ領域144の各ページは、主記憶領域管理テーブル142のエントリと1対1に対応している。同様に、拡張記憶装置102上の各ページは、拡張記憶領域管理テーブル143のエントリと1対1に対応している。

【0031】アドレス変換テーブル141、主記憶領域管理テーブル142、拡張記憶領域管理テーブル143は、それぞれ主記憶装置101内に存在する。

【0032】図5にアドレス変換テーブル141の構造を示す。アドレス変換テーブル141の各エントリは、実記憶アドレス151、拡張記憶アドレス152、磁気ディスクアドレス153および拡張記憶割当済フラグ154の各フィールドを有する。を有する。

【0033】図6に主記憶領域管理テーブル142の構造を示す。主記憶領域管理テーブル142の各エントリは、仮想アドレス155およびページ属性フラグ157の各フィールドを有する。

【0034】図7に拡張記憶領域管理テーブル143の構造を示す。拡張記憶領域管理テーブル143の各エントリは、仮想アドレス155および再利用可能フラグ156の各フィールドを有する。図5～図7の各テーブルは相互に参照できる。本実施例では、主記憶アドレスおよび拡張記憶アドレスのかわりに、テーブルのエントリのアドレスや番号を用いている。主記憶領域管理テーブル142のページ属性157は、アドレス変換テーブル141上にあってもよい。

【0035】図5に示したアドレス変換テーブル141の拡張記憶アドレス152および拡張記憶割当て済フラグ154は、仮想記憶の領域が拡張記憶装置102が割当てられたときセットされる。マイグレーションされたとき、すなわち拡張記憶装置102からその領域が磁気ディスク装置103へ追い出されたとき、拡張記憶割当て済フラグ154はリセットされるが、拡張記憶アドレス152はそのまま残される。これにより、拡張記憶装置102の領域がマイグレーションされても、その領域が他の仮想記憶に割当てられるまでの間に同じ仮想記憶領域についてページインが起こると、その領域を再利用し、I/Oを削減することができる。ある仮想記憶領域のアドレス変換テーブルの拡張記憶アドレス152に対応したESTEが、その仮想記憶領域のアドレスを保持していれば拡張記憶装置102の当該領域は再利用可能

である。

【0036】テーブルの各エントリはページ単位に設けられるため、テーブルの領域として多くの主記憶領域が使われる。そこで、主記憶領域を節約するため、アドレス変換テーブル141の拡張記憶アドレス152のフィールドを他の用途に利用する場合がある。例えば、磁気ディスク装置103を二重化した場合、二重化した磁気ディスク装置のアドレスと拡張記憶アドレス152とで同じフィールドを共有する。また、仮想入出力のバッファでは、磁気ディスクアドレスの代わりにファイル中の位置とファイルの識別番号とをアドレス変換テーブル141に格納する必要がある。この場合には、拡張記憶アドレス152のフィールドをファイルの識別番号の格納に利用する。

【0037】拡張記憶アドレス152のフィールドを他の用途に利用した場合、拡張記憶装置102が割当てられたときは拡張記憶アドレス152を保持しているが、マイグレーションされると例えば磁気ディスクアドレス153を保持する。このようにマイグレーションされた場合は拡張記憶アドレス152を失ってしまうため、アドレス変換テーブル141から以前割当てられていた拡張記憶領域を求めることができず、拡張記憶領域の再利用は不可能である。

【0038】したがって、マイグレーションの対象となる領域として、再利用可能な領域を選択した方が、再度拡張記憶装置102へページインする際に処理が簡略化され、好ましい。再利用可能な領域のみを選択してマイグレーションするためには、再利用可能な領域と不可能な領域を、マイグレーションに先だって判断する必要がある。そこで、拡張記憶領域管理テーブル143 (図7)の各エントリに、そのエントリに対応した領域に格納するデータが、拡張記憶アドレス152の領域を他の用途に利用しているか否かを示す再利用可能フラグ156をもたせる。再利用可能フラグ156は、拡張記憶装置102へページアウトするとき、二重化または仮想入出力などのページの属性を示すフラグ157 (図6)を参照しセットする。

【0039】再利用可能フラグ156がセットされていなければ、拡張記憶装置102の空き領域に余裕があるマイグレーションの場合、マイグレーションの対象から外す。これにより再利用可能な領域のみをマイグレーションできるため、マイグレーションを開始する空き容量 (第2基準空き容量) を高く設定しても、磁気ディスク装置103からのページインの増加を防ぐことができる。

【0040】次に、仮想記憶制御装置105の各制御部の処理について説明する。まず、マイグレーション制御部114の処理について説明する。

【0041】図8に、マイグレーション処理のフローチャートを示す。マイグレーションは、ページアウト処理

の完了時、または資源集中管理プログラムにより一定間隔をおいて起動される。起動時の空き容量により、2段階の処理に分かれる。このマイグレーション処理を説明する前に、この処理に用いる拡張記憶制御テーブル160の構成を図9により説明する。拡張記憶制御テーブル161は、図1には示さなかったが、拡張記憶領域管理部155内に設けられる。

【0042】図9に示すように、拡張記憶制御テーブル160は、マイグレーションを開始する空き容量の基準値として、第1基準空き容量161と、第2基準空き容量162の2つを持つ。その他、拡張記憶装置のページングデータ格納容量163、拡張記憶装置の空き容量164、使用中キュー123の末尾ポインタ166、マイグレーション開始ポインタ121、使用中キューの先頭ポインタ165、利用可能キュー124の末尾ポインタ168、マイグレーション開放領域先頭ポインタ122、利用可能キュー124の先頭ポインタ167を持つ。この拡張記憶制御テーブル160の各ポインタと拡張記憶領域管理テーブル143（図7）の各エントリ（131～135）との関係も図9に示す。

【0043】さて、図8のマイグレーション処理において、まず、空き容量163が第1基準空き容量161以下の場合は（ステップ301）、使用中領域のキュー123の先頭のESTEに対応した領域132cから（すなわちページアウトされた時刻の古い順に）、第1レベル空き容量と現在の空き容量との差の容量分だけ、拡張記憶装置から主記憶を経由して磁気ディスク装置へデータを転送する（ステップ302）。転送後、当該領域が再利用可能領域であれば（ステップ303）、利用可能キューの末尾ポインタ168の示すESTE133aの次に末尾の領域として追加し、拡張記憶装置102の当該領域を解放する（ステップ304）。また、再利用不可能な領域のESTEは、利用可能キュー124の先頭ポインタ167の示すESTE135aの前に先頭の領域として追加する（ステップ305）。これにより、この先頭の領域は、ページインにより解放された領域134a、134bより優先して他の仮想記憶のページに割当てられることになる。

【0044】空き容量が第1基準空き容量161以上第2基準空き容量162以下の場合は（ステップ306）、一定の起動間隔をおいて一定量のマイグレーションを行う。前回のマイグレーション起動から起動間隔に達していないときは、マイグレーションを行わない（ステップ307）。起動間隔は、数秒程度を想定しているが、マイグレーション処理のCPUオーバーヘッドと空き容量の変動割合を考慮し設定する。1回の起動によりマイグレーションされる量は、ファイルへのI/Oを妨げないようにするため、ページングI/Oによるチャネルおよび磁気ディスクの使用率を100%未満とする範囲で設定する（ステップ308）。また、1回の起動によ

りマイグレーションされる量は、空き容量に応じて変更してもよい。

【0045】マイグレーション開始ポインタ121の指すESTE131bから使用中キュー123を順に（図では右から左に）たどって再利用可能か否かを調べ、再利用可能な領域であれば磁気ディスク装置103に転送する（ステップ309）。なお、使用中キュー123の末尾には、後述するページアウト処理により新たなESTEが追加されていく。磁気ディスク装置103への転送後、当該ESTEを利用可能キュー124の末尾ポインタ168の示すESTE133aの次に新たな末尾ESTEとして追加し、転送した拡張記憶領域を解放する（ステップ310）。再利用不可能な領域であれば、ESTEはそのまま使用可能なキュー123に残す。マイグレーション完了後は、調べ終ったESTEの次のESTEにマイグレーション開始ポインタ121を移動する（ステップ311）。これにより、マイグレーション開始ポインタ121の前（図では右側）には、再利用不可能な領域のESTEのみが残る。

【0046】次に、ページアウト制御部112の処理について説明する。図10に、本実施例のページアウト処理のフローチャートを示す。

【0047】拡張記憶装置102に空き領域があるかどうかを調べ（ステップ401）、空き領域がなければ磁気ディスク装置103に主記憶の内容を転送する（ステップ402）。空き領域があれば、利用可能キュー124の先頭のESTE135aに対応した領域を割当てて、そのESTEを使用中キュー123の末尾ポインタ166の示すESTE133aの次に移動し（ステップ403）、割当てた領域に主記憶の内容を転送する（ステップ404）。このとき、ページインにより解放された領域を管理するESTEが、マイグレーションにより解放された領域を管理するESTEよりキューの前方に位置するため、利用可能キュー124の先頭のESTEの示す拡張記憶領域を割当てることにより、マイグレーションによって解放された領域よりページインにより解放された領域が優先して割り当てられる。ページインによって解放された拡張記憶領域は、その拡張記憶領域が他の仮想記憶領域に割当てられても、先に当該拡張記憶領域に存在したページは主記憶装置に取り込まれており、そのページへのアクセスは何ら問題なく、そのページのページアウト時も、CPU命令により高速に実行される。しかし、マイグレーションによって解放された拡張記憶領域が他の仮想記憶領域に割当てられると、マイグレーションにより追い出されたページは磁気ディスク装置にのみ存在することになるので、当該ページへのアクセス時にはI/O命令により主記憶装置へ転送しなければならない。そのため、マイグレーションによって解放された拡張記憶領域を拡張記憶装置に残すことにより、ページングI/Oを削減することができる。

【0048】また、拡張記憶アドレス152のフィールドを他の用途に利用しているか否かをアドレス変換テーブル141のページ判別フラグ157等により調べ（ステップ405）、利用していなければ再利用可能フラグ156をセットする（ステップ406）。転送後は、主記憶領域を解放する（ステップ407）。

【0049】次に、ページイン制御部111の処理について説明する。図11に、本実施例のページイン処理のフローチャートを示す。

【0050】主記憶領域を仮想記憶領域に割当てたのち（ステップ501）、ページインするデータが拡張記憶装置102に存在するかどうか、すなわち、拡張記憶装置102を割当て済（使用中キュー123内）または再利用可能（利用可能なキュー124内）であるかどうか、を調べる（ステップ502）。存在しなければ磁気ディスクからページインする（ステップ503）。存在すれば、当該ページのデータを拡張記憶装置102から主記憶装置101に転送後（ステップ504）、転送した拡張記憶領域に対応したESTEを使用中のキューから利用可能キュー124のマイグレーション解放領域ポインタ122が示すESTE（図7の133b）の前の位置に移動する（ステップ505）。この操作により、利用可能キュー124において、ページインにより解放された領域のESTE134aおよび134bは、マイグレーションにより解放された領域のESTE133a、133bの前方に配置される。以降、拡張記憶領域を割当てる場合は、利用可能キュー124の先頭のESTEに対応した領域を割当てることにより、ページインにより解放された領域をマイグレーションにより解放された領域より優先して割当てることができる。

【0051】また、ページインする拡張記憶領域を管理するESTEをマイグレーション開放領域ポインタ122が指していれば、換言すれば、マイグレーション解放領域ポインタ122が指しているESTEの領域のデータがページインされた場合（ステップ506）、マイグレーション開放領域ポインタ122を次のマイグレーションにより解放された拡張記憶領域を管理するESTEに移動する（ステップ507）。

【0052】最後に、スワップ制御部113の処理について説明する。図12に、本発明のスワップ処理のフローチャートを示す。

【0053】拡張記憶装置102の容量が十分存在し、スワップアウトされたデータがマイグレーションされるまでの時間が十分確保される場合は（ステップ601）、メモリ不足を原因としてスワップアウトされたデータの殆どがマイグレーションされる前にスワップインされることが期待できる。このような場合では、スワップアウトする原因にかかわらず、拡張記憶装置102に空き領域があれば（ステップ602）拡張記憶装置へスワップアウトし（ステップ603）、空き領域がない場

合のみ磁気ディスク装置にスワップアウトする（ステップ604）。

【0054】一方、拡張記憶装置の容量が十分でない場合は、資源集中管理から起動時に入力パラメタとして受け取ったスワップアウトの原因を調べる（ステップ605）。メモリ不足を原因としてスワップアウトされるデータであれば、拡張記憶装置に割当て可能であっても、磁気ディスク装置へ出力する（ステップ606）。端末入出力待ちなどその他の原因であれば、拡張記憶装置102に空き領域があれば拡張記憶装置へスワップアウトする（603）。

【0055】拡張記憶の容量が十分であるか否かは、以下の基準により判断する。メモリ不足を原因としてスワップアウトされたデータがスワップインされるまでの平均時間が、拡張記憶装置に出力されてからマイグレーションされるまでの時間より短ければ、拡張記憶装置からスワップインされることが期待できる。拡張記憶装置に出力されてからマイグレーションされるまでの時間は、拡張記憶容量／ページアウト頻度で近似できる。従って、メモリ不足を原因としてスワップアウトされたデータがスワップインされるまでの平均時間と、拡張記憶装置へのページアウト頻度を測定し、拡張記憶容量が、（メモリ不足を原因としてスワップアウトされたデータがスワップインされるまでの平均時間）＊平均ページアウト頻度より少なければ、拡張記憶装置からスワップインされることが期待できないので磁気ディスクに出力し、他のデータの拡張記憶装置からの入力の可能性を増やす。

【0056】

【発明の効果】以上のように、本発明によれば、負荷の変動が激しい場合において高速外部記憶装置を有効利用し高速外部記憶装置の容量を増やさずとも高負荷時の高速外部記憶装置の空き領域の供給不足を解消できるとともに、低速外部記憶装置からの転送を削減することができ、仮想記憶制御方式の計算機システムの性能を向上させることができる。

【図面の簡単な説明】

【図1】本発明の仮想記憶制御システムの構成を示すブロック図

【図2】従来の仮想記憶制御システムの構成を示すブロック図

【図3】従来のマイグレーション制御方式を示したフローチャート

【図4】本発明の仮想記憶制御システムのハードウェアイメージ図

【図5】アドレス変換テーブルの構成を示す図

【図6】主記憶管理テーブルの構成を示す図

【図7】拡張記憶管理テーブルの構成を示す図

【図8】本発明のマイグレーション制御方式を示したフローチャート

【図9】拡張記憶制御テーブルおよび拡張記憶領域管理テーブルの関連図

【図10】本発明のページアウト制御方式を示したフローチャート

【図11】本発明のページイン制御方式を示したフローチャート

【図12】本発明のスワップ制御方式を示したフローチャート

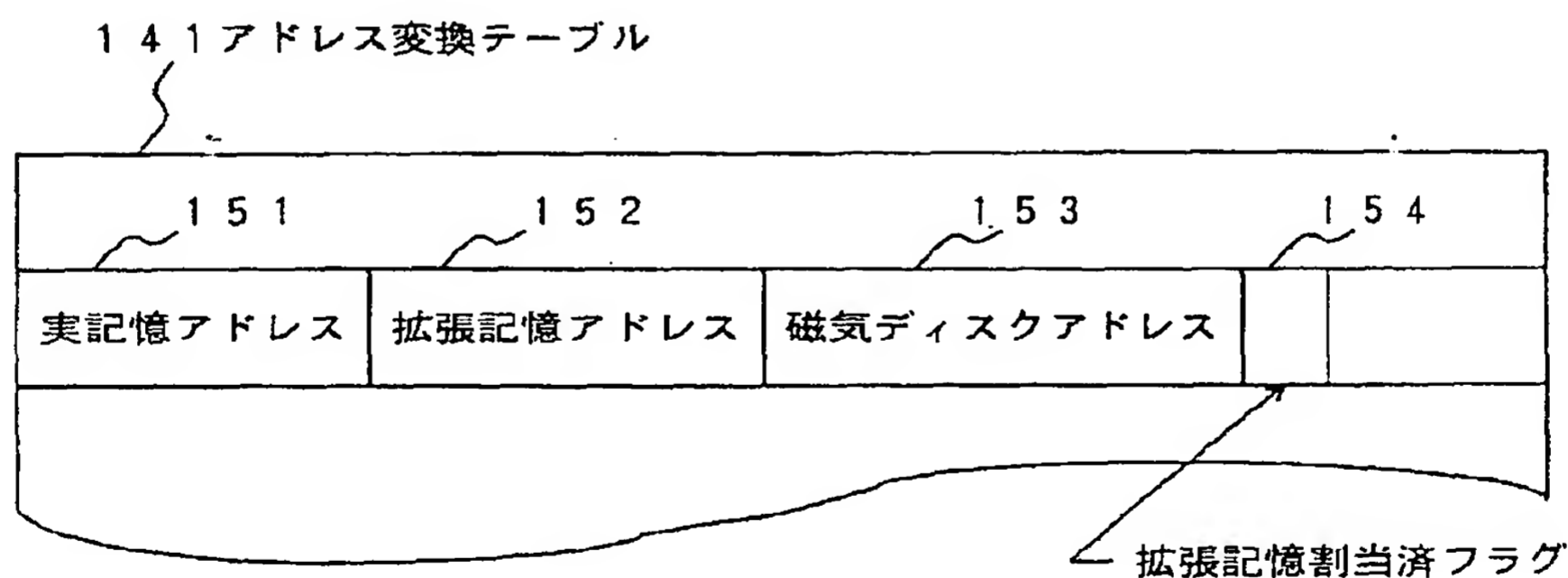
【符号の説明】

101…主記憶装置、102…拡張記憶装置、103…磁気ディスク装置、104…仮想記憶装置、105…仮想記憶制御装置、106…CPU（中央処理装置）、107…IOP（入出力処理装置）、111…拡張記憶領域管理部、112…ページイン制御部、113…ページアウト制御部、114…スワッピング制御部、115…マイグレーション制御部、116…スワップ原因判別部、117…転送量調整部、118…転送拡張記憶領域選択部、121…マイグレーション開始ポインタ、122…マイグレーション解放領域ポインタ、123…使用中の拡張記憶領域を管理するテーブルのキュー、124…利用可能な拡張記憶領域を管理するテーブルのキュー、131a、131b、131c、131d、131e…使用中の再利用可能な拡張記憶領域を管理するテ

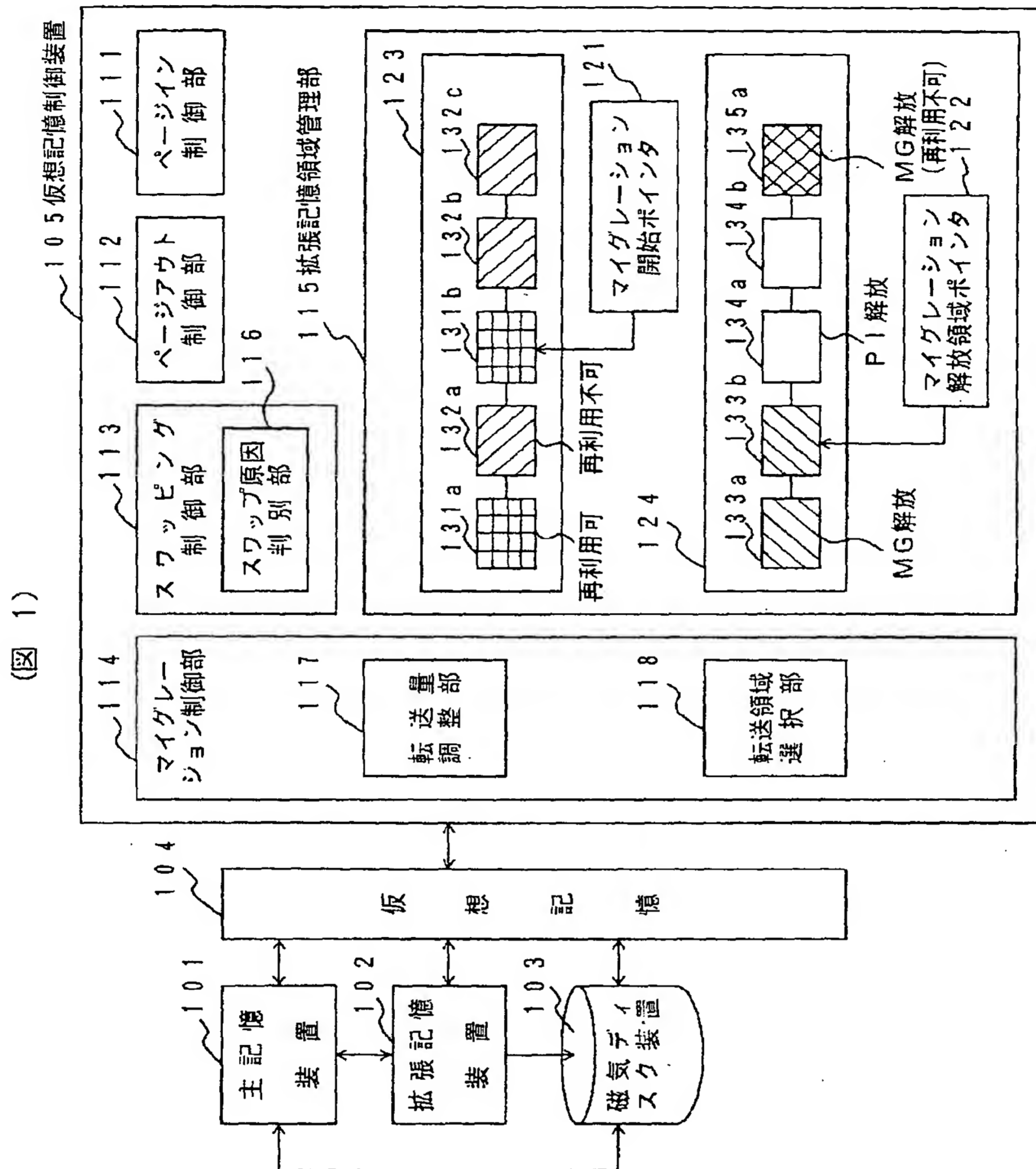
ブル、132a、132b、132c、132d…使用中の再利用不可能な拡張記憶領域を管理するテーブル、133a、133b、133c、133d…マイグレーションにより解放された再利用可能な拡張記憶領域を管理するテーブル、134a、134b、134c、134d…ページインにより解放された拡張記憶領域を管理するテーブル、135a、135b…マイグレーションにより解放された再利用不可能な拡張記憶領域を管理するテーブル、141…アドレス変換テーブル、142…主記憶管理テーブル、143…拡張記憶管理テーブル、144…主記憶中のデータ格納領域、151…実記憶アドレス、152…拡張記憶アドレス、153…磁気ディスクアドレス、154…拡張記憶割当て済フラグ、155…仮想記憶アドレス、156…再利用可能フラグ、157…ページ属性フラグ、160…拡張記憶制御テーブル、161…第1基準空き容量、162…第2基準空き容量、163…拡張記憶装置のページングデータ格納容量、164…拡張記憶装置の空き容量、165…使用中キューの先頭の拡張記憶管理テーブルを示すポインタ、166…使用中キューの末尾の拡張記憶管理テーブルを示すポインタ、167…利用可能キューの先頭の拡張記憶管理テーブルを示すポインタ、168…利用可能キューの末尾の拡張記憶管理テーブルを示すポインタ。

【図5】

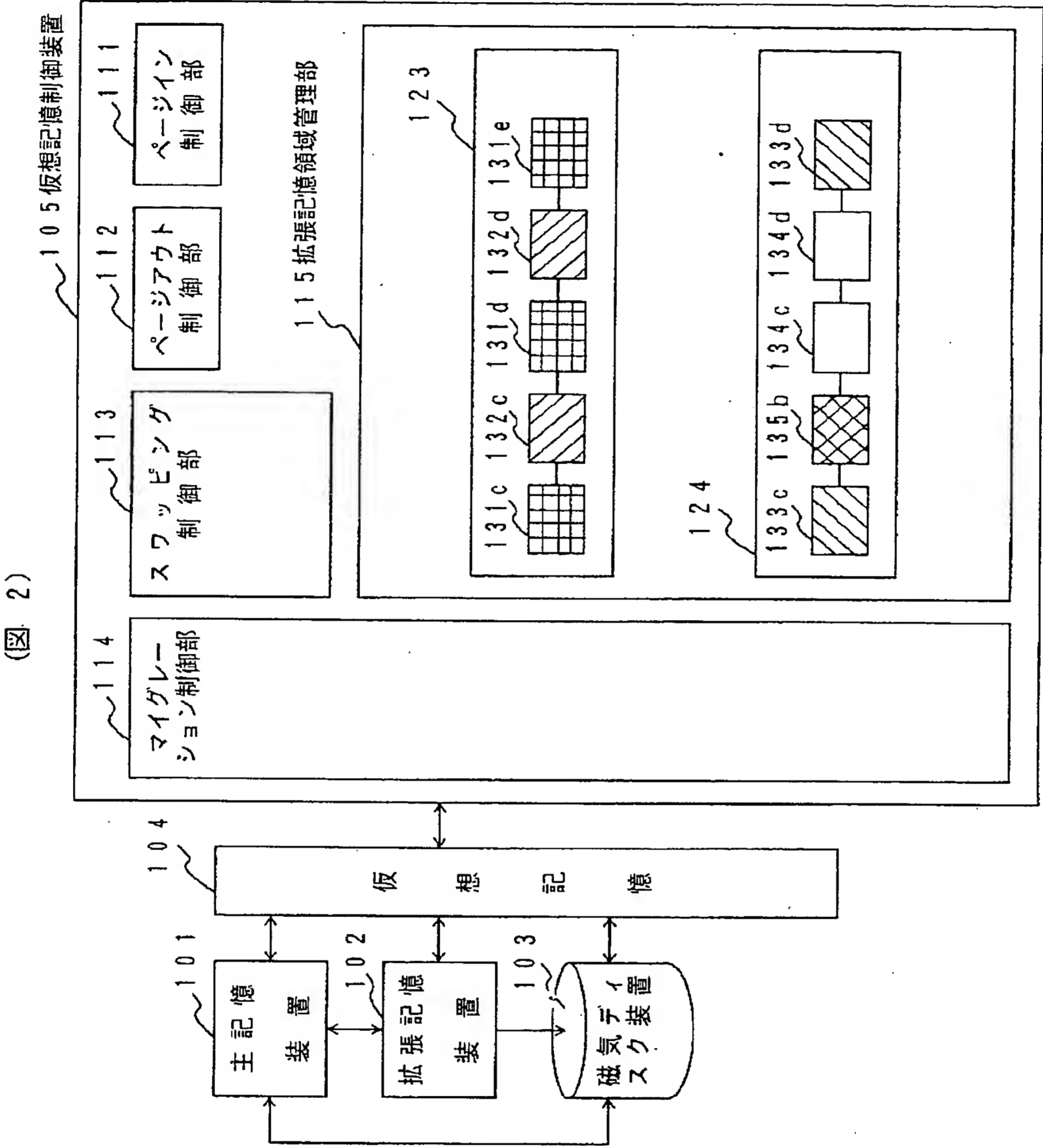
(図 5)



【図1】

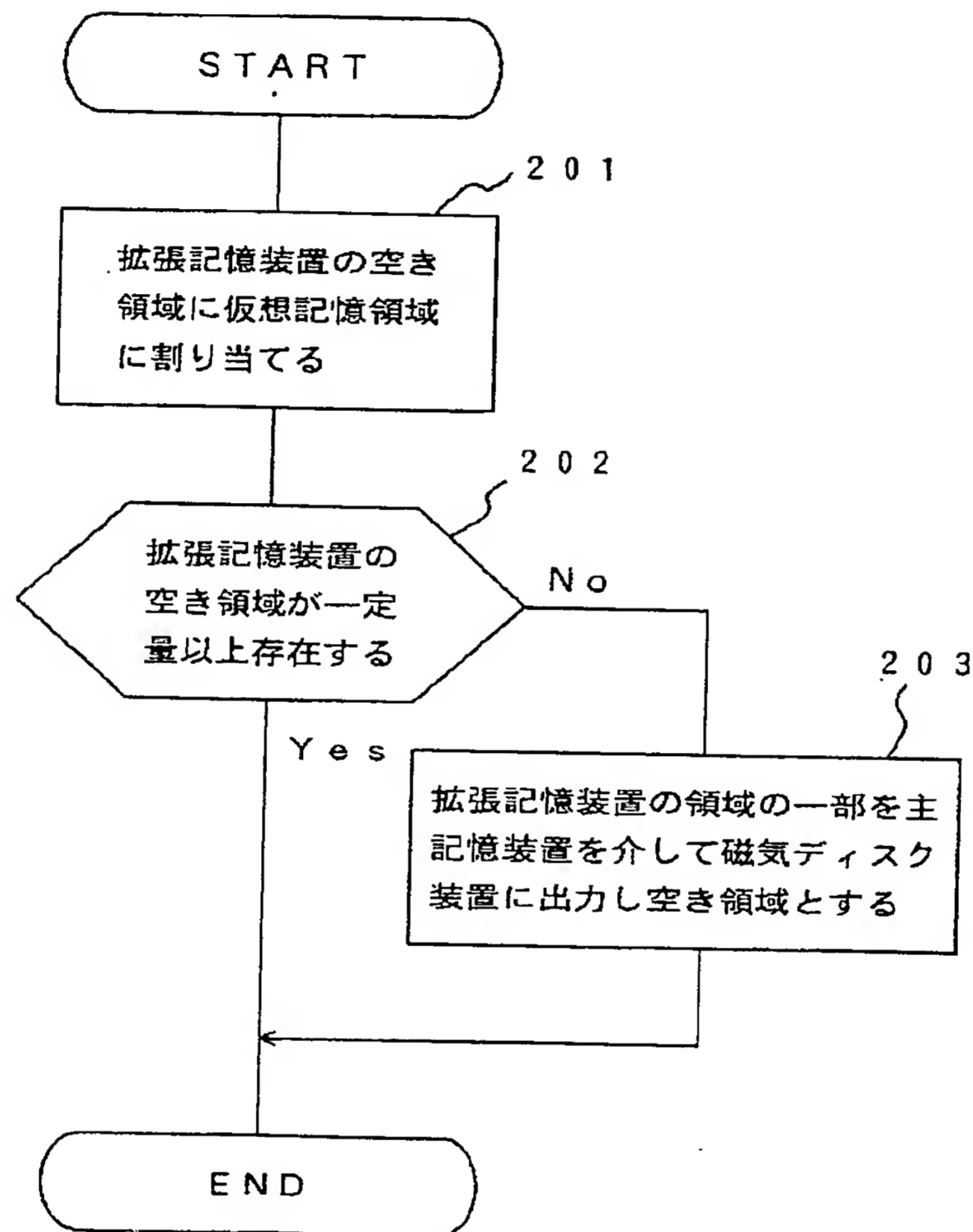


【図2】

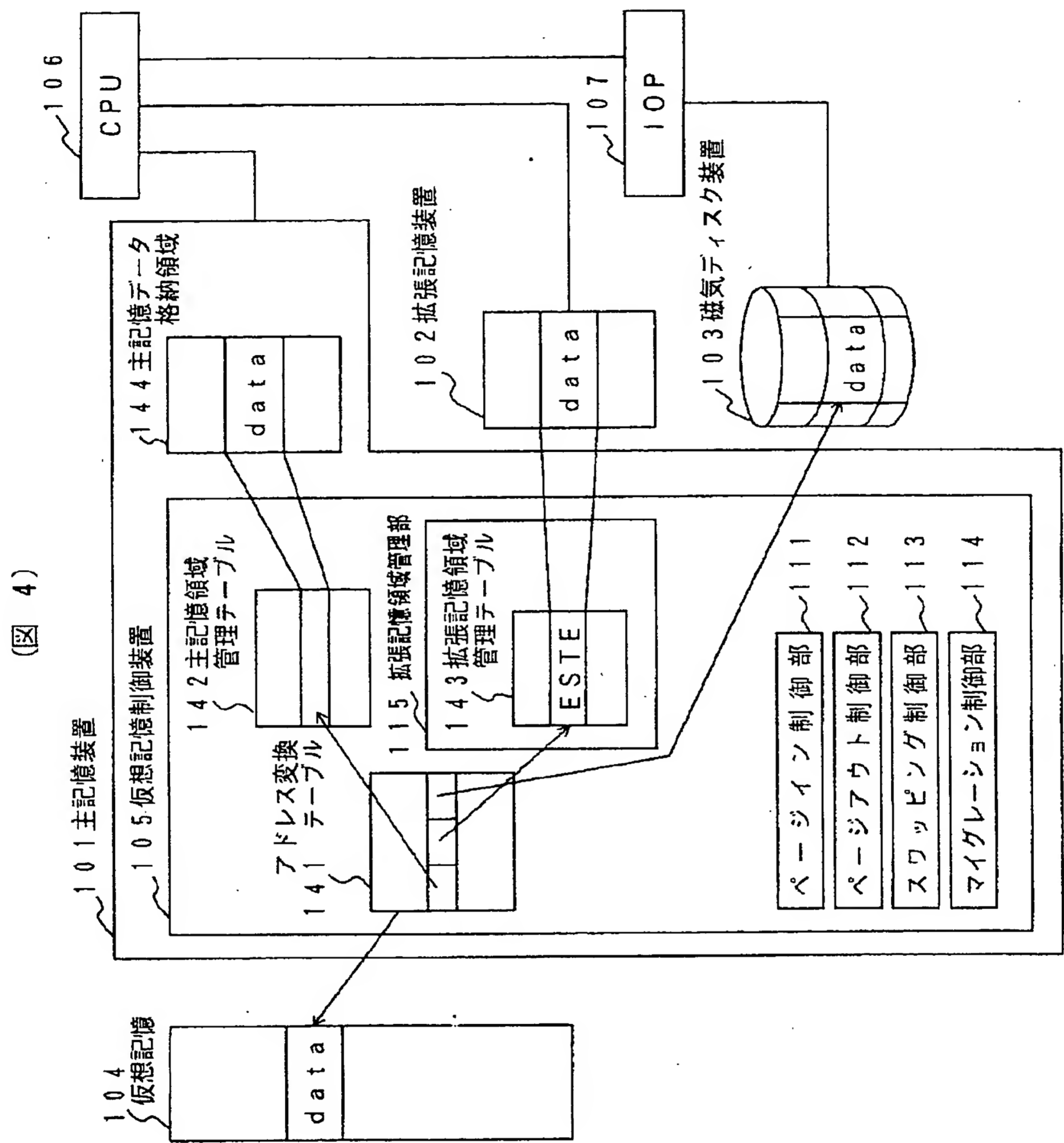


【図3】

(図 3)



【図4】



【図6】

(図 6)

142 主記憶領域管理テーブル

155			157 ページ属性フラグ	
仮想アドレス				

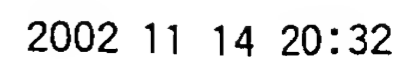
【図7】

(図 7)

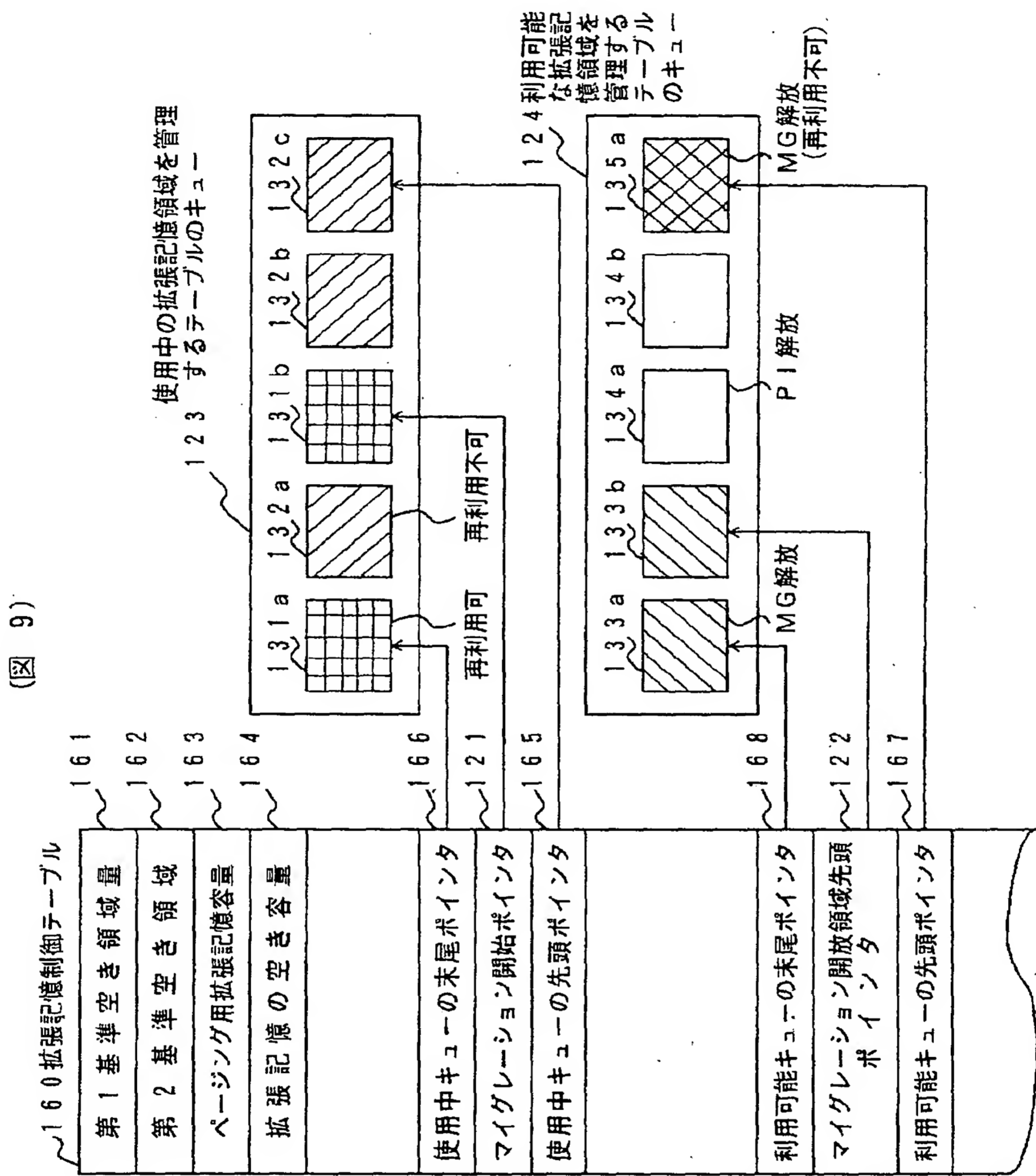
143 拡張記憶領域管理テーブル

155		
仮想アドレス		
156 再利用可能フラグ		

マイグレーション処理 (図 8)

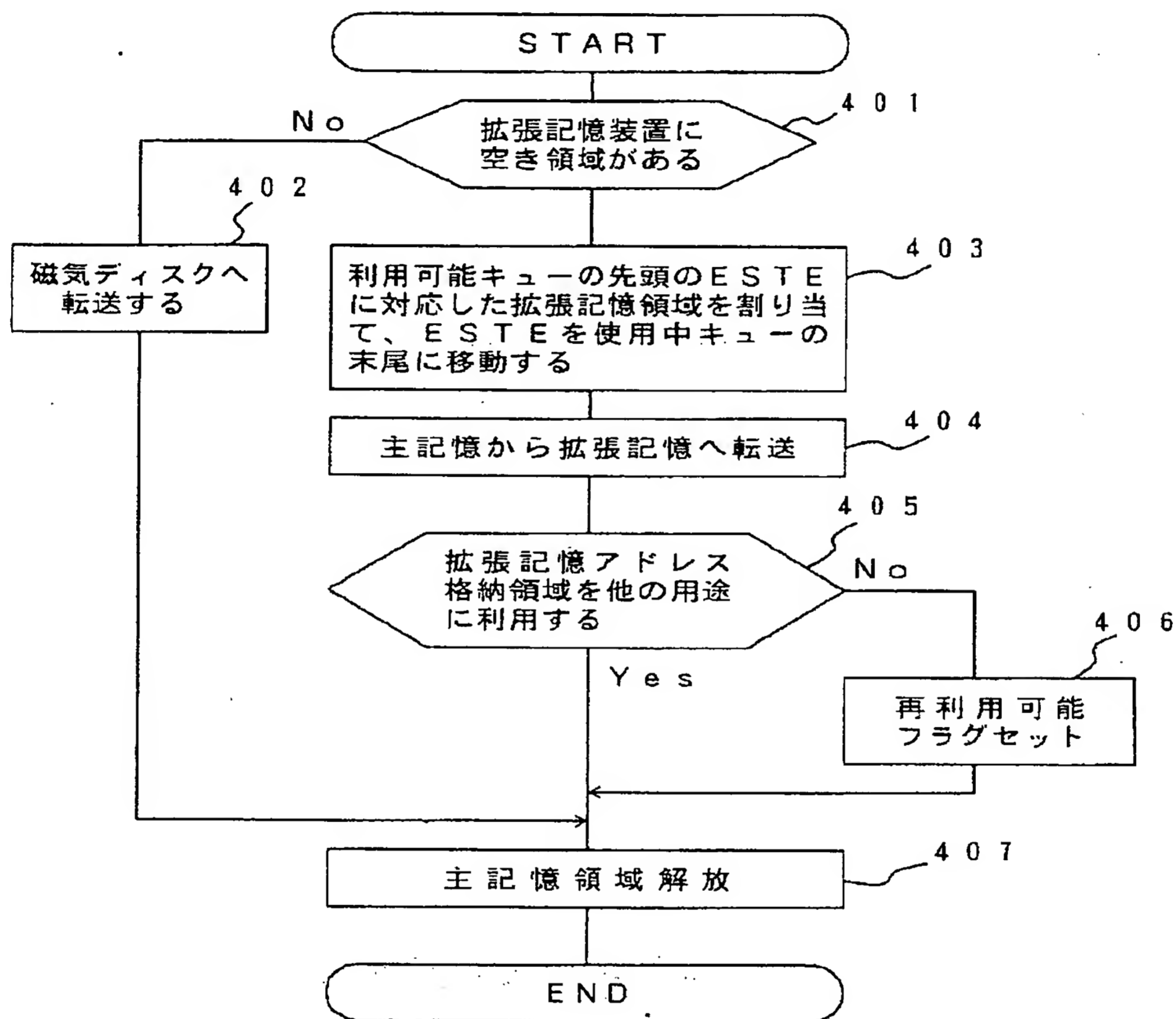


【図9】



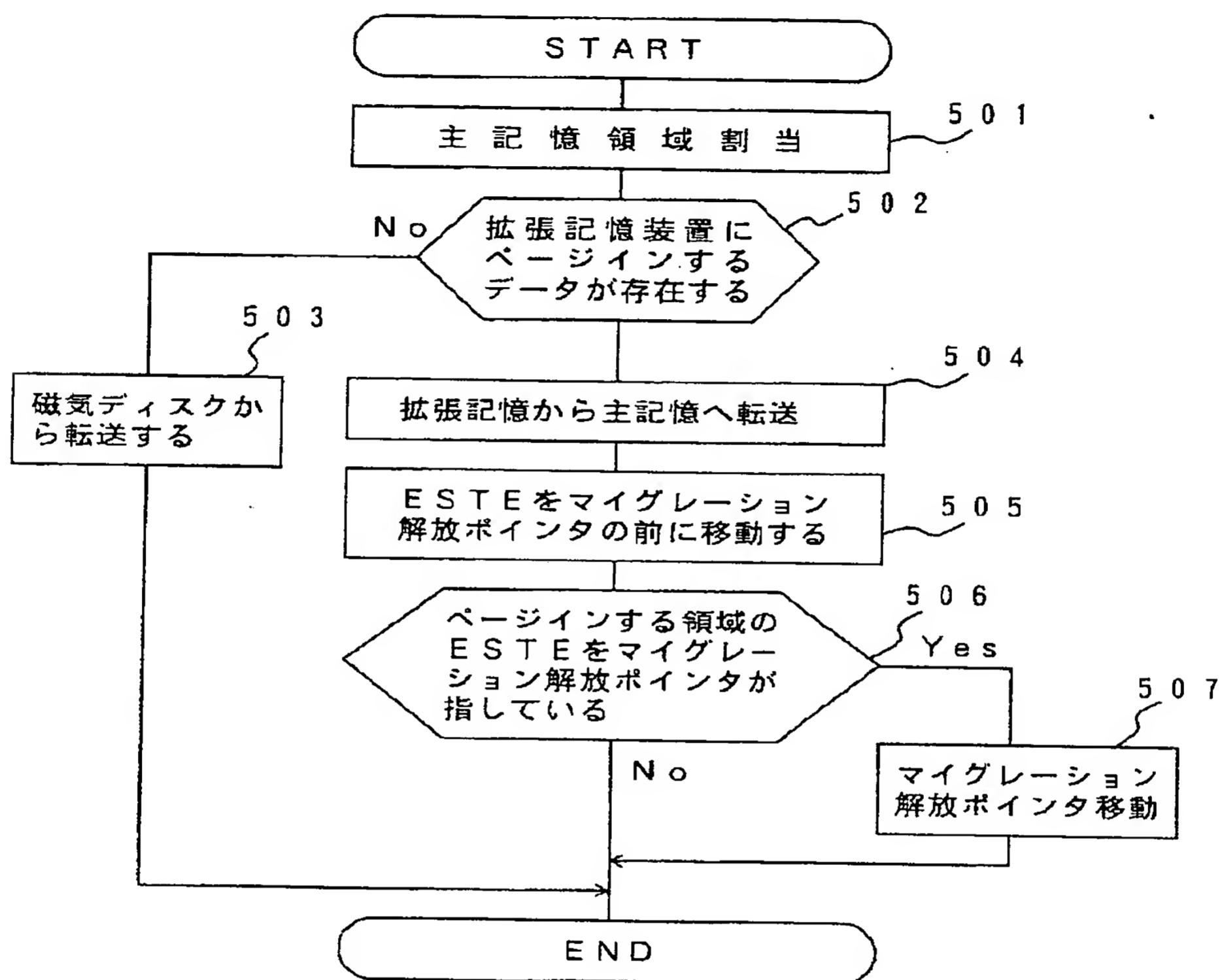
【図10】

ページアウト処理 (図 10)



【図11】

ページイン処理 (図 11)



【図 12】

スワップ処理 (図 12)

